Protein Science (1995), *4*:1161–1166. Cambridge University Press. Printed in the USA.

Incorporation of pairwise interactions into the Lifson–Roig model for helix prediction

WILLIAM SHALONGO AND EARLE STELLWAGEN
Department of Biochemistry, University of Iowa, Iowa City, Iowa 52242

(RECEIVED November 21, 1994; ACCEPTED February 10, 1995)

Abstract

The helix/coil equilibrium of a peptide in solution can be modulated by a variety of side-chain interactions that are not incorporated into the standard statistical mechanical models for prediction of peptide helical content. In this report, we describe a recursive formulation of the Lifson–Roig model that facilitates incorporation of specific pairwise side-chain interactions as well as nonspecific individual side-chain capping interactions. Application of this extended model to a series of host/guest peptides indicates that the apparent ΔG value for a pairwise apolar interaction is dependent upon the spacing and orientation but not the sequential location of the participating residues. The apparent ΔG values for such interactions are about 40% greater than the corresponding apparent $\Delta\Delta G$ values obtained from difference measurements.

Keywords: apolar interactions; capping interactions; helical peptides; helix prediction; Lifson–Roig model; side-chain interactions

Although an individual residue in a monomeric helical peptide can be legitimately analyzed in terms of a simple two-state helix/coil equilibrium, the peptide as a whole cannot be so analyzed. This pertains because the helix/coil equilibrium constant of a residue within the frayed ends of a peptide helix is diminished relative to the helix/coil equilibrium constant for a residue in the central region of a helix. Such diminution results from the loss of backbone hydrogen bonding and the increased entropy of a residue in the frayed ends. Accordingly, any measure of the mean helical content of a peptide solution, such as the mean residue ellipticity at 222 nm, cannot be rigorously analyzed using a simple two-state helix/coil equilibrium.

However, measurements of the mean helical content of peptide solutions can be rigorously analyzed using statistical mechanical models that express the overall helix/coil equilibrium of a peptide in terms of the individual helix/coil equilibrium for each constituent residue. In order to express this relationship, all statistical mechanical models make four common assumptions, namely: (1) a residue in a given peptide molecule can only populate a helical state or a nonhelical (coil) state; (2) the helical state of each residue is the same; (3) the fractional helical content of a given residue ranges from 0 to 1; and (4) a residue in the coil state is given a statistical weight of 1 if its adjacent residues are also in the coil state. By contrast, individual statistical mechanical models are distinguished as to (1) what feature of a residue, its α-carbon, peptide bonds, or ϕ, ψ angles, is considered in the assignment of states; (2) whether nonamino acid backbone moieties such as terminal blocking groups are considered residues; (3) how the statistical weights of a residue are affected by the states of its adjacent residues; and (4) how interactions modulate helix/coil equilibria.

The widely used Zimm–Bragg (1959) and Lifson–Roig (1961) statistical mechanical models do not consider such interactions. Recently, Doig et al. (1994) have extended the Lifson–Roig model by reweighing some coil states to simulate unspecified capping interactions generated by individual residues. However, this extended model, denoted here as the capping model, does not recognize pairwise interactions involving specific residues located in a helical conformation. In this report, we describe an alternative extension of the Lifson–Roig model, denoted here as the interaction model, which can accommodate specific pairwise interactions as well as unspecified capping interactions.

Description of the models

All the statistical mechanical models to be described calculate the ensemble of coil, partially helical, and helical peptide states as a canonical partition function, $Z(r)$, where r is the number of residues (α carbons) in the peptide. Acetyl and amide blocking groups, when present, are counted as residues for calculation of $Z(r)$. The helical and coil states of an individual residue, i, are statistically weighted according to the states of its adjacent residues. These statistical weights, v_i, w_i, n_i, and c_i, are essentially equilibrium constants for a given combination of states

Reprint requests to: Earle Stellwagen, Department of Biochemistry, University of Iowa, Iowa City, Iowa 52242; e-mail: cmdste@vaxa.weeg.uiowa.edu.

and can be expressed as free energies. The statistical weights employed by each of the three models used in this study are summarized in Table 1.

The Lifson–Roig model calculates the fractional helical content of the ith residue of a peptide, $f_{\alpha,i}$, using Equation 1:

$$f_{\alpha,i} = \partial \ln Z(r)/\partial \ln w_i. \quad (1)$$

The statistical weight v_i for each residue and the statistical weight w_i for each nonterminal residue contribute to the partition function $Z(r)$. Statistical weights n_i and c_i do not contribute to Equation 1. Because the statistical weight w_i for each terminal residue does not contribute to Equation 1, the terminal residues are perpetually in coil states. Although each blocking group contributes a statistical weight v_i to $Z(r)$, the blocking groups are excluded from the residue count when calculating the mean helical content of the entire peptide.

The capping model reweighs three of the coil states, as shown in Table 1, to account for capping interactions. These three coil states were selected to maintain the terminal residues and the capping residue in perpetual coil states. The capping model calculates the fractional helical content of the ith residue of a peptide using Equation 2:

$$f_{\alpha,i} = \partial \ln Z(r)/\partial \ln w_i + \partial \ln Z(r)/\partial \ln v_i. \quad (2)$$

This expansion of Equation 1 is necessary so that the reweighed coil states can affect the helical content of the entire peptide. As in the Lifson–Roig model, the blocking groups are excluded from the residue count when calculating the mean helical content of the entire peptide.

The interaction model requires recursive equations rather than the traditional matrix equations to include pairwise interactions between the side chains of specific residues. The interaction model calculates the fractional helical content of the ith residue of a peptide using Equation 3:

$$f_{\alpha,i} = Z'(r)_i/Z(r). \quad (3)$$

The term $Z'(r)_i$ is the derivative of the partition function $Z(r)$ for the ith residue. Values for $Z(1)$ through $Z(r)$ and $Z'(1)_i$ through $Z'(r)_i$ are calculated by incrementing j, a counter denoting the current residue being considered. Equations 4 and 5,

$$Z(j) = \sum_{k=-1}^{j-1} Z(k)F(j, k+2) \quad (4)$$

$$Z'(j)_i = \sum_{k=-1}^{j-1} Z'(k)_i F'(j, k+2)_i, \quad (5)$$

describe $Z(r)$ and $Z'(r)_i$, respectively, for residues 1 through j; $Z(j)$ and $Z'(j)_i$, in terms of the partition functions already calculated for residues 1 through k; and $Z(k)$ and $Z'(k)_i$, where $k \leq j$. The initial partition functions are special conditions, namely $Z'(i-1)_i$ is 0 and $Z(-1)$, $Z(0)$, $Z'(-1)_i$, and $Z'(0)$ are each 1. The terms $F(j, k+2)$ and $F'(j, k+2)_i$ are the statistical weights for an all-helical segment spanning residues j through $k+2$. Values for these statistical weights are listed in Table 2 under the column labeled "No interactions." Because each terminal residue can have a finite helical content (as discussed below), terminal blocking groups must be included in the residue count when calculating the mean helical content of the entire peptide.

The recursive equations can only be extended to include capping interactions by selectively changing the statistical weights of two helical states of the Lifson–Roig model as indicated in Table 1. These two new statistical weights, vn and vc, appear in the expressions for $F(j, k+2)$ and $F'(j, k+2)_i$ when $k+2$ is equal to either $j-1$ or $\leq j-2$, as listed in Table 2 under the columns headed "Capping interactions" and "Capping and pairwise interactions." These changes in the statistical weights of the helical states allow the terminal residue in a peptide to have a finite helical content. Recent ^{13}C chemical shift measurements (Shalongo et al., 1994) suggest this to be the case.

Each pairwise interaction is introduced into the recursive equations in terms of a donor residue, d, and an acceptor residue, a. These two residues as well as all the intervening residues must be in the helical state. The fractional helical content of each of these residues is modulated by the pairwise interaction expressed as an apparent ΔG. Equation 6,

$$K_m(j, k+2) = e^{-\Delta G/RT} \text{ if } (k+2 \leq d) \text{ and } (j \geq a), \quad (6)$$

expresses the ΔG for the pairwise interaction in terms of an equilibrium constant, $K_m(j, k+2)$. If the limits of Equation 6 do not pertain, $K_m(j, k+2)$ is assigned a value of 1. Multiple pairwise interactions involving residues in the span j to $k+2$ contribute to a collective equilibrium constant, $E(j, k+2)$, as described in Equation 7:

$$E(j, k+2) = \prod_{m=1}^{p} K_m(j, k+2). \quad (7)$$

The collective equilibrium constant appears in the expression for $F(j, k+2)$ and $F'(j, k+2)_i$ when $k+2$ is equal to j, $j-1$, or $\leq j-2$, as listed in Table 2 under the columns headed "Capping interactions" and "Capping and pairwise interactions."

Use of recursive Equations 4–7 for evaluation of $Z(j)$ for the first two residues in a peptide containing both capping and pairwise interactions is illustrated in Table 3. In considering this illustration, it should be remembered that j is the residue number;

Table 1. *Assignment of statistical weights in each model*[a]

Residue states			Statistical weights		
$i-1$	i	$i+1$	Lifson–Roig	Capping	Interaction
C	**C**	C	1	1	1
C	**C**	H	1	n	1
H	**C**	C	1	c	1
H	**C**	H	1	$\sqrt{(nc)}$	1
C	**H**	C	v	v	v
C	**H**	H	v	v	vn
H	**H**	C	v	v	vc
H	**H**	H	w	w	w

[a] The state of the residue contributing the statistical weight to the partition function $Z(r)$ is shown in boldface type.

Table 2. *Evaluation of $F(j, k+2)$ and $F'(j, k+2)_i$*

	$F(j, k+2)$ or $F'(j, k+2)_i$			
$k+2$	No interactions	Capping interactions	Pairwise interactions	Capping and pairwise interactions
$>j+1$	0	0	0	0
$j+1$	1	1	1	1
j	v_{k+2}	v_{k+2}	$E(j, k+2)v_{k+2}$	$E(j, k+2)v_{k+2}$
$j-1$	$v_j v_{k+2}$	$(vc)_j(vn)_{k+2}$	$E(j, k+2)(v)_j(v)_{k+2}$	$E(j, k+2)(vc)_j(vn)_{k+2}$
$\leq j-2$	$v_j v_{k+2} \prod_{m=j-1}^{k+3} w_m$	$(vc)_j(vn)_{k+2} \prod_{m=j-1}^{k+3} w_m$	$E(j, k+2)v_j v_{k+2} \prod_{m=j-1}^{k+3} w_m$	$E(j, k+2)(vc)_j(vn)_{k+2} \prod_{m=j-1}^{k+3} w_m$

k ranges from -1 to $j-1$, as stated in Equation 4; $Z(k)$ is $Z(j)$ calculated for the previous residue, except for the special conditions for $Z(-1)$ and $Z(0)$ stated in the text; expressions for $k+2$ in terms of j and the corresponding value for $F(j, k+2)$ are given in Table 2; and $Z(j)$ is defined in Equation 4.

The Lifson–Roig, capping, and interaction models have been written as menu-driven FORTRAN programs for a personal computer. The programs require the sequence and either the mean helical content of no more than 115 peptides or the residue helical contents of no more than 30 peptides. Statistical weights and interaction-apparent ΔG values are refined using the Brent (1973) algorithm to avoid difficulties in calculating the partial derivatives of the fitting functions. The parameters to be fit during each iteration are selected in a different order to minimize any bias due to parameter selection.

Refinement of statistical weights in the absence of pairwise interactions

Statistical weights in the family of Lifson–Roig models do not describe physical properties defining the helix state, such as the probability that a residue has helical ϕ,ψ angles or that a residue makes one or more backbone hydrogen bonds. The values of the statistical weights are essentially arbitrary model-dependent parameters defined by the mathematical relationships built into the model. Therefore, each model will generate a different set of legitimate statistical weights when fit against a common database comprising the mean fractional helical content of a series of peptides of defined sequence. The goodness of fit between the observed and predicted mean helical contents of all the peptides in a database is expressed as an RMS deviation.

The mean fractional helical content of 57 peptides containing apolar and uncharged polar guest residues has served previously as a database for refinement of statistical weights (Chakrabartty et al., 1994). The sequences of these peptides, listed in Table 4, were designed to minimize pairwise side-chain interactions. Measurements were obtained using solvent conditions that maximize helical content and minimize electrostatic interactions among host residues. In keeping with precedent (Chakrabartty et al., 1994; Doig et al., 1994), each model was fit with this database at a series of fixed v weights allowing the w and n weights to float freely. All n weights are relative to an n weight of 1.0 for alanine, a residue assumed to be devoid of N-terminal capping interactions (Doig et al., 1994). All c weights are assumed to be 1.0 because the peptide sequences in the database were designed primarily to investigate N-terminal capping interactions (Doig et al., 1994). No specific pairwise interactions were used in the interaction model to fit the 57 peptide database.

Our programming of the capping model generates an RMS of 3.1, as reported by Chakrabartty et al. (1994), using a common v weight of 0.048 and the w and n weights listed in their Tables 2 and 3. However, the minimum RMS generated by the capping model, 2.72, occurs at a common v weight of 0.034, as shown in Figure 1. Residue w and n weights obtained by the capping model at this minimal RMS are presented in Table 5.

If the acetyl and amide blocking groups are included in the residue count for the predicted mean helical content of each peptide, the capping model exhibits an RMS minimum, 2.91, at a common v weight of 0.034. The residue w weights obtained are changed by a mean value of 0.02 (0.02) with the number in parentheses indicating the standard deviation. The n weights for the majority of the residues are diminished 0.20 (0.05). The n

Table 3. *The initiation of a recursive calculation*[a]

j	k	$Z(k)$	$k+2$	$F(j, k+2)$	$[Z(k)][F(j, k+2)]$
1	−1	1	1	$E(1, 1)v_1$	$[1][E(1, 1)v_1]$
1	0	1	2	1	[1][1]
2	−1	1	1	$E(2, 1)(vn)_1(vc)_2$	$[1][E(2, 1)(vn)_1(vc)_2]$
2	0	1	2	$E(2, 2)v_2$	$[1][E(2, 2)v_2]$
2	1	$[E(1, 1)v_1] + 1$	3	1	$[[E(1, 1)v_1] + 1][1]$

[a] $Z(1) = [E(1, 1)v_1] + 1$ and $Z(2) = E(1, 1)v_1 + E(2, 1)(vn)_1(vc)_2 + E(2, 2)v_2 + 1$.

Table 4. *Observed and predicted mean fractional helical contents*[a]

Peptide sequence	Mean fractional helical content		
		Predicted using	
	Observed	Capping model	Interaction model
NAKAAAAKAAAAKAAGYm	0.60	0.58	0.60
GAKAAAAKAAAAKAAGYm	0.53	0.47	0.50
SAKAAAAKAAAAKAAGYm	0.53	0.53	0.53
TAKAAAAKAAAAKAAGYm	0.46	0.45	0.46
LAKAAAAKAAAAKAAGYm	0.45	0.43	0.43
IAKAAAAKAAAAKAAGYm	0.42	0.42	0.42
MAKAAAAKAAAAKAAGYm	0.40	0.39	0.40
PAKAAAAKAAAAKAAGYm	0.40	0.40	0.40
AAKAAAAKAAAAKAAGYm	0.38	0.37	0.37
VAKAAAAKAAAAKAAGYm	0.38	0.37	0.37
QAKAAAAKAAAAKAAGYm	0.29	0.30	0.30
KALAAKALAAKAAAAKGGYm	0.29	0.33	0.32
KAMAAKAMAAKAAAAKGGYm	0.24	0.26	0.25
KAQAAKAQAAKAAAAKGGYm	0.22	0.22	0.21
KAAAAKAIAAKAAAAKGGYm	0.29	0.24	0.25
KAAAAKASAAKAAAAKGGYm	0.21	0.23	0.24
KAAAAKANAAKAAAAKGGYm	0.21	0.22	0.22
KAAAAKAVAAKAAAAKGGYm	0.20	0.19	0.20
KAAAAKATAAKAAAAKGGYm	0.18	0.17	0.17
aKAAAAKAWAAKAAAAKm	0.50	0.50	0.50
aKAAAAKAYAAKAAAAKm	0.51	0.51	0.51
aKALAAKALAAKAAAAKGGYm	0.49	0.52	0.51
aKAMAAKAMAAKAAAAKGGYm	0.40	0.43	0.43
aKAQAAKAQAAKAAAAKGGYm	0.40	0.43	0.42
aKAAAAKAIAAKAAAAKGGYm	0.48	0.49	0.48
aKAAAAKASAAKAAAAKGGYm	0.45	0.45	0.45
aKAAAAKANAAKAAAAKGGYm	0.43	0.41	0.41
aKAAAAKAVAAKAAAAKGGYm	0.43	0.41	0.41
aKAAAAKATAAKAAAAKGGYm	0.40	0.35	0.35
aYGAAKAAAAKAAAAKAA	0.49	0.51	0.52
aYGAAKAAAAKAAAAKAG	0.48	0.51	0.52
aYGKAAAAKAAAAKAAAAKm	0.74	0.66	0.65
aYGGKAAAAKAAAAKAAAAKm	0.68	0.64	0.62
aYGGKAAAAKALAAKAAAAKm	0.55	0.58	0.57
aYGGKAAGAKAAAAKAAAAKm	0.37	0.34	0.36
aYGGKAAAAKAAGAKAAAAKm	0.26	0.27	0.29
aYGGKAAAAKAAAAKAAGAKm	0.47	0.48	0.41
aYGGKAAAAKAMAAKAAAAKm	0.51	0.53	0.53
aYGGKAAAAKAQAAKAAAAKm	0.51	0.54	0.53
aYGGKAAAAKAIAAKAAAAKm	0.46	0.48	0.48
aYGGKAAAAKACAAKAAAAKm	0.45	0.45	0.45
aYGGKAAAAKASAAKAAAAKm	0.43	0.44	0.45
aYGGKAAAAKAFAAKAAAAKm	0.43	0.43	0.43
aYGGKAAAAKANAAKAAAAKm	0.40	0.40	0.41
aYGGKAAAAKATAAKAAAAKm	0.30	0.34	0.36
aYGGKALAAKALAAKALAAKm	0.49	0.44	0.45
aYGGKAMAAKAMAAKAMAAKm	0.36	0.32	0.33
aYGGKAQAAKAQAAKAQAAKm	0.31	0.32	0.32
aYGGKASAAKASAAKASAAKm	0.22	0.20	0.19
aYGGKAIAAKAIAAKAIAAKm	0.22	0.22	0.22
aYGGKANAAKANAAKANAAKm	0.16	0.18	0.16
aYGGKAVAAKAVAAKAVAAKm	0.09	0.13	0.11
aYGGKATAAKATAAKATAAKm	0.08	0.11	0.07
aYGGQAAAAQAAAAQAAAAQm	0.52	0.53	0.52
aYGGQAAAAQAQAAQAAAAQm	0.38	0.40	0.41
aYGGQAQAAQAQAAQAQAAQm	0.27	0.21	0.22
aYGGGKAAAAKAAAAKAAAAKm	0.62	0.61	0.59

[a] Observed mean fractional helical contents of these peptides were reported by Chakrabartty et al. (1994). Mean fractional helical contents were predicted using the capping model and the interaction model with capping but not pairwise interactions. The residue *v*, *w*, and *n* weights used by each model in these predictions are listed in Table 5.

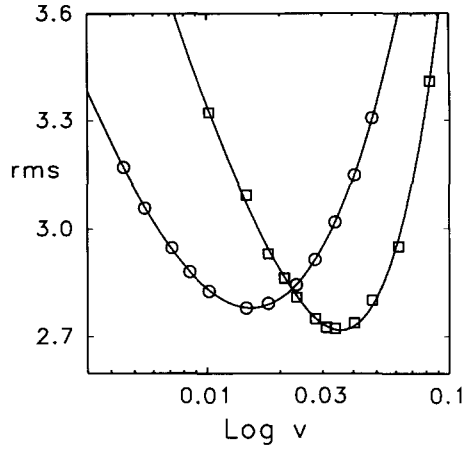

Fig. 1. Dependence of the RMS deviation between the observed and predicted mean fractional helical contents of the peptides listed in Table 4 on the *v* statistical weight. Values obtained using the interaction model (○) and the capping model (□), respectively, for refinement are presented. Values for the *w* and *n* statistical weights obtained at the minimum for each model are listed in Table 5.

weights of glycine, asparagine, and the acetyl group are increased 0.8 (0.2) and the *n* weight of glutamine is decreased by 0.6. These comparisons indicate that use of the residue count employed in the interaction model only introduces modest changes in the parameters generated by the capping model.

Table 5. *Refined statistical weights values*[a]

Residue	*w* weights		*n* weights	
	Capping model	Interaction model	Capping model	Interaction model
A	1.67	1.77	1.0	1.0
C	0.42	0.58		
F	0.36	0.50		
G	0.11	0.22	3.7	3.0
I	0.51	0.76	2.1	1.6
K	1.03	1.17	0.6	0.3
L	1.01	1.19	2.4	1.8
M	0.72	0.92	1.5	1.3
N	0.38	0.47	11.5	5.9
P	<0.001	<0.001	1.7	1.4
Q	0.69	0.93	0.1	0.4
S	0.45	0.60	6.4	3.7
T	0.21	0.33	3.0	2.2
V	0.30	0.48	1.1	1.0
W	0.53	0.60		
Y	0.57	0.62	3.0	1.4
Acetyl			4.8	3.3

[a] These statistical weight values were obtained at the *v* weight for each model that gives the minimal RMS value, $v = 0.034$ for the capping model and $v = 0.015$ for the interaction model, as shown in Figure 1. These weights were refined in the absence of pairwise interactions between specific residues. The absence of values for some statistical weights indicates that they could not be obtained from the peptide sequences listed in Table 4.

The interaction model exhibits a minimal RMS of 2.78 at a v weight of 0.015, as shown in Figure 1. This RMS value is comparable to the minimal RMS value obtained using the capping model. The small v weight at the RMS minimum for the interaction model increases the residue w weights and decreases the residue n weights compared with the capping model, as shown in Table 5. The interdependence of residue n, v, and w weights for selected residues in the interaction model is illustrated in Figure 2. This interdependence suggests that the differences in the w and n weights generated by the capping model and by the interaction model, shown in Table 5, are principally the result of a different v weights in the RMS minimum found by each model. As shown in Figure 2, all residue w weights and all residue n weights greater than 0.2 exhibit a linear dependence on the logarithm of their v weights. The high population of short peptide helices that exists at high v weights forces the small n weights to become even more deleterious to helix stability.

Evaluation of pairwise interactions

Pairwise interactions between specific residue side chains that affect helix/coil equilibria include, among others, ion pairs, hydrogen bonds, and apolar interactions. Apparent $\Delta\Delta G$ values associated with particular pairwise interactions have been estimated from measurements of the mean helical contents of host/guest peptides in variable solvent conditions. Such analyses are flawed because the mean helical content of peptide solutions cannot be analyzed rigorously using a simple two-state analysis, and because helical peptide states are populated that do not contain the given interaction, resulting in an underestimate of the apparent ΔG for the given interaction.

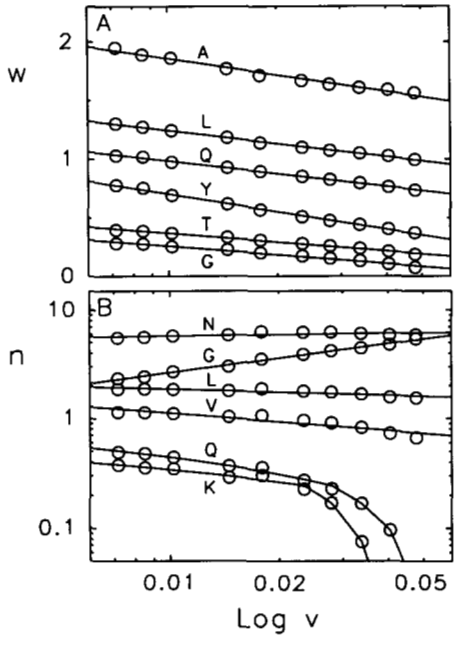

Fig. 2. Dependence of the w **(A)** and n **(B)** statistical weights of selected residues on the fixed v statistical weight using the interaction model. These values were obtained from analysis of the peptides listed in Table 4, assuming the absence of pairwise interactions.

Padmanabhan and Baldwin (1994) have recently described a series of peptides, each containing a single potential apolar sidechain interaction. The fractional mean helical content of each of these peptides was measured using the same solvent conditions employed in the measurement of the fractional helical contents of the peptides listed in Table 4. Accordingly, the statistical weights listed in Table 5 should be appropriate for prediction of the observed helical contents of the peptides described by Padmanabhan and Baldwin (1994). In the absence of pairwise interactions, the interaction model underpredicts the observed helical content of all but one of the peptides, as shown in Table 6. In the presence of pairwise interactions, the interaction model can predict the apparent ΔG of the single apolar interaction in each peptide, which would equate the observed and predicted helical content of that peptide. These apparent ΔG values are listed in Table 6.

Two residues involved in a potential pairwise interaction are denoted here as X/Z with residue X being closer to the N-terminus. The mean apparent ΔG for the Y/V and Y/L i, $i+3$ interactions is -0.06 kcal/mol, as shown in Table 6. This value suggests that apolar i, $i+3$ interactions are, at best, very weak as surmised (Padmanabhan & Baldwin, 1994). By contrast, the mean apparent ΔG for the Y/V and Y/L i, $i+4$ interactions is significantly larger, -0.58 kcal/mol. The magnitude of this value suggests that i, $i+4$ apolar interactions make a significant contribution to the stability of monomeric peptide helices.

As shown in Table 6, the apparent ΔG values for both the i, $i+3$ and i, $i+4$ L/Y interactions are significantly enhanced relative to the corresponding values for the Y/L interactions. This enhancement likely reflects the different side-chain rotamer populations of tyrosine and leucine in the helix state, trans and gauche, respectively (Dunbrack & Karplus, 1993; Schrauber et al., 1994). These rotamers would generate a stabilizing i, $i+3$ interaction when oriented L/Y but not Y/L. In contrast, these rotamers would generate a stabilizing i, $i+4$ interaction when oriented either L/Y or Y/L. These results suggest that both the spacing and orientation of pairwise apolar residues is important to peptide helix stability.

As shown in Table 6, the apparent ΔG predicted for each significant pairwise apolar interaction is larger than its apparent $\Delta\Delta G$ by a factor of 1.4 (0.2). Recognition of such a factor is important to estimation of ΔG values from $\Delta\Delta G$ measurements.

Alternative analyses

The statistical weights described in this report predict the helix/coil equilibria of a larger number of peptides with good fidelity. Regrettably, these statistical weights cannot be related to either geometric or energetic features of the peptides. Unfortunately, molecular dynamic and ab initio calculations, which consider these features, are not designed to model helix/coil equilibria. These calculations are best suited to model and refine a limited collection of related peptide structures with similar conformational energies. Construction of an entire ensemble of peptide structures at equilibrium would require a great deal of computer time and a multiplicity of arbitrary choices. Accordingly, statistical mechanical models, which are explicitly designed to model helix/coil equilibria, appear to represent the current method of choice for prediction of the fractional helical content of peptide solutions.

Table 6. *Analysis of pairwise apolar interactions*[a]

Sequence	Fractional helical content		ΔG (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
	Observed	Predicted		
$i, i + 3$ interactions				
a**Y**K**AV**AAKAAAAKAAAAKm	0.53	0.47	−0.24	−0.13
a**Y**KA**L**AAKAAAAKAAAAKm	0.59	0.62	0.12	0.07
aAA**Y**KA**L**AAKAAAm	0.20	0.16	−0.17	−0.15
aAKAAA**Y**KA**L**AAKAAAAKm	0.52	0.53	0.07	0.02
Mean (SD)			−0.06 (0.18)	
$i, i + 4$ interactions				
a**Y**KAA**V**AKAAAAKAAAAKm	0.59	0.45	−0.49	−0.31
a**Y**KAA**L**AKAAAAKAAAAKm	0.75	0.61	−0.61	−0.36
aAA**Y**KAA**L**AKAAAm	0.36	0.16	−0.69	−0.59
aAKAAA**Y**KAA**L**AKAAAAKm	0.71	0.55	−0.53	−0.36
Mean (SD)			−0.58 (0.08)	
Reverse orientations				
aAAAKAA**L**AK**Y**AAm	0.33	0.18	−0.58	−0.44
aAAAKA**L**AAK**Y**AAm	0.42	0.18	−0.84	−0.65

[a] Residues in boldface type indicate the only potential pairwise apolar interaction considered in each peptide. Observed fractional helical contents of each peptide were measured by Padmanabhan and Baldwin (1994). Predicted fractional helical content for each peptide was obtained using the interaction model, the statistical weight values v, w, and n listed for this model in Table 4, and NO pairwise interactions. ΔG value for each peptide indicates the stability of its constituent pairwise apolar interaction, which makes the observed and predicted helical contents identical. $\Delta\Delta G$ values for each peptide represent the difference between the observed and predicted helical contents expressed as free energies, $\Delta\Delta G = (RT \ln f_{\alpha,\text{pred}}) - (RT \ln f_{\alpha,\text{obs}})$.

Acknowledgments

This work was supported in part by National Research Service Award GM15759 (to W.S.) from the National Institute of General Medical Sciences and by the Department of Biochemistry and the College of Medicine at the University of Iowa.

References

Brent RP. 1973. *Algorithms for minimization without derivatives*. Englewood Cliffs, New Jersey: Prentice Hall. pp 61–80.

Chakrabartty A, Kortemme T, Baldwin RL. 1994. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci 3*:843–852.

Doig AJ, Chakrabartty A, Klingler TM, Baldwin RL. 1994. Determination of free energies of N-capping in α-helices by modification of the Lifson–Roig helix-coil theory to include N- and C-capping. *Biochemistry 33*: 3396–3403.

Dunbrack RL Jr, Karplus M. 1993. Backbone-dependent rotamer library for proteins; application to side-chain prediction. *J Mol Biol 230*:543–574.

Lifson S, Roig A. 1961. On the theory of helix-coil transition in polypeptides. *J Chem Phys 34*:1963–1974.

Padmanabhan S, Baldwin RL. 1994. Helix-stabilizing interaction between tyrosine and leucine or valine when the spacing is $i, i + 4$. *J Mol Biol 241*:706–713.

Schrauber H, Eisenhaber F, Argos P. 1994. Rotamers: To be or not to be? An analysis of amino acid side-chain conformations in globular proteins. *J Mol Biol 230*:592–612.

Shalongo W, Dugad L, Stellwagen E. 1994. Analysis of the thermal transitions of a model helical peptide using ^{13}C NMR. *J Am Chem Soc 116*: 2500–2507.

Zimm BH, Bragg JK. 1959. Theory of phase transition between helix and random coil in polypeptide chains. *J Chem Phys 31*:526–535.